Defense Technical Information Center
Compilation Part Notice

ADP012060

TITLE: A Comparative Study of Expansion Functions Using the Boundary Residual Method on a Linear Dipole - Part II: Sub-Domain Functions

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Applied Computational Electromagnetics Society Journal. Volume 17, Number 1

To order the complete compilation report, use: ADA399939

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP012055 thru ADP012066

A Comparative Study of Expansion Functions Using the Boundary Residual Method on a Linear Dipole – Part II: Sub-Domain Functions

Malcolm M. Bibby
mbibby@gullwings.com

Abstract. The Boundary Residual Method, BRM, is used to compare the performance of a range of sub-domain expansion functions in three different situations. It is found that as the complexity of the structure/system being studied increases the importance of the order of the expansion function decreases – at least for uniform segmentation. These results point out the need for more research into the behavior and performance of these, and other, expansion functions.

Introduction.

In a companion paper [1], published in the same issue of this journal, a wide range of entire-domain expansion functions was identified and those functions were compared for convergence – in both a global and a local sense. The intent of this work is to provide similar information for a wide range of sub-domain functions. There are many more sub-domain candidates than entire-domain candidates. Consequently, only a sub-set of the possibilities, albeit a significant number, will be examined.

Sub-domain expansion functions were introduced for two main reasons. First, there was the intent to make the computation of the associated integro-differential equations easier. Second, it was recognized that entire-domain expansion functions were difficult, if not impossible, to use on curves/surfaces that contained discontinuities. Early sub-domain expansion functions were simple – pulses, and triangles. Subsequently, more geometrically sophisticated functions were proposed – ones that addressed the perceived shortcomings of the simpler forms. These shortcomings centered mainly on the continuity, or more precisely the lack of it, at the intercepts of the expansion functions. The newer functions were also intended to conform to the surfaces that they were applied to. Unfortunately, these developments proceeded without a thorough comparison of the performance of the new function with that of the prior art. Consequently, today there is only sparse understanding of the merits and drawbacks of each of the alternatives and in particular which ones are "better". [Better will be defined and discussed later]. In 1975, Butler and Wilton [2] examined various numerical techniques which included a study of some of the simpler expansion functions applied to a dipole. Their work also involved examination of the form of the integro-differential equation and the testing function and results were presented in terms of the current at the center of the dipole. No estimates of global convergence were reported and the local convergence rates were only reported visually and not quantitatively. Today, much improved tools are available to overcome these deficiencies. With this in mind, many of the more widely used sub-domain expansion functions were examined in a consistent manner in different applications.

In this paper, the various relevant mathematical issues are first reviewed. The subsequent section contains the results of examining the various expansion functions and discusses those findings. Finally, a number of conclusions are presented and suggestions for further work made.

Methodology.

Error Minimization. The present numerical work relies on least squares minimization in the manner discussed by Bunch and Grow [3][4]. This approach was reviewed in Part I of this two part series and is not repeated here. It is important to note that, using this method, one obtains an explicit measure of the residual error resulting from the expansion function employed. This measure permits one to track the global convergence – that is, how well the expansion functions perform over the entire surface and not just at one position such as at the center of a dipole.

In anticipation of the study of spline expansion functions, some additional description of Least Squares Estimation, LSE, is necessary [5][6]. In particular, it is useful to consider the least squares solution when the variables are required to satisfy

1054-4887 © 2002 ACES

specified linear equality constraints. In such an instance, we wish to solve, in a least squares sense, the equation $\vec{M}A = \vec{B}$ subject to $CA = D$. $\vec{M}$ is an $(m \times n)$ matrix, A is an $(n \times 1)$ vector representing the solution and $\vec{B}$ is an $(m \times 1)$ vector representing the excitation. C is a $(p \times n)$ matrix and D is a $(p \times 1)$ vector, subject to $m > n > p$. [In the case of spline functions there are continuity conditions at certain knots that must be satisfied. These p conditions provide the input for C and D]. This conditional minimization procedure is described elsewhere [5][6] and code is available [7].

Convergence. Most text/research books in the CEM area refer to errors decreasing as $O(h^p)$ where p is the degree of the polynomial underlying the basis/expansion function used. The $O(h^p)$ model can be developed into a formal equation. Specifically, consider when one has computed data for an observable, Y_n, using n expansion functions and one wishes to extrapolate to an asymptotic value, Y_∞. In the case of a single-dimensional problem, the appropriate model is:

$$Y_n = Y_\infty + O(h^p) = Y_\infty + \beta n^\alpha. \tag{1}$$

One expects $\alpha = -p$. The 'outer' equation can be differentiated and combined with the original to produce

$$Y_n = Y_\infty + \frac{dY_n}{dn}.\frac{n}{\alpha} = Y_\infty + \frac{1}{\alpha}\left(n\frac{dY_n}{dn}\right). \tag{2}$$

This equation permits estimations of both Y_∞ and α through standard linear regression techniques. These regression techniques are well developed and allow one to determine the reliability of the estimates of the two constants. Here, $\frac{dY_n}{dn}$ is replaced by $\frac{\Delta Y_n}{\Delta n}$. A plot of Y_n versus $\frac{\Delta Y}{\Delta n}.n$, where Y_n represents the current for a given number of expansion functions, indicates that this is a very good model in many cases and can be used to extrapolate for Y_n. Unfortunately, although widely available linear regression tools can be used to solve (2) for Y_n and α, they cannot be used to make predictive statements about the reliability of the values so obtained. Most linear regression tools expect the independent variable to be essentially error free. Such a claim cannot be made for $\frac{\Delta Y}{\Delta n}$ and this makes the reliability claims questionable. Nevertheless, the plots can be made and visually examined to see whether the model is applicable. On the other hand, if we use (1) and recast it as:

$$\log(Y_n - Y_\infty) = \log(\beta) + \alpha\log(n) \tag{3}$$

then we can use conventional linear regression. This formulation requires prior knowledge of Y_∞, typically the term of most interest and in such instances only provides information about α. Equation (3) is used in this study.

Numerical Procedures.

General. The numerical procedures used here were the same as in the companion paper [1] with the following exceptions

1) The SVD routine was replaced by the appropriate routine for constrained least squares, ZGGLSE, from LAPACK 3.0 [7].
2) The value of m, in the matrices, was set at $3(N+q)$. This multiple was found by numerical experimentation to provide reliable, stable values. A minimum multiple of 2 was suggested by Ikuno and Yasuura [9].
3) The term $\frac{\Delta X}{\Delta n}$ was computed using a simple central difference approach.

Edge Mode. It is well-known that whenever surfaces contain physical discontinuities they require special treatment. For example, most practical structures, with the notable exception of complete loops, incorporate edges. Nevertheless few, if any, sub-domain functions incorporate the edge mode/condition into their formulation. A model that does not incorporate consideration of the edge mode is inherently deficient. In the current work, the edge mode is incorporated explicitly. This will be illustrated in the context of a straight dipole. According to theory developed by Meixner [10], the current at the edge of a dipole with an infinitely thin wall thickness is proportional to d^ν where ν=0.5 and d is the distance from the end of the dipole. In this work, the end segments are $\left(\frac{y-h}{\Delta y_0}\right)^\nu$ and $\left(\frac{(h-y)}{\Delta y_{n+1}}\right)^\nu$ where y is the

Table I. A description of the eight expansion functions used in the current study.
$1 \le i \le n$

Legend	Type	Mathematical Description
P	Pulse	$I(y) = \begin{cases} 1 & y_i < y < y_{i+1} \\ 0 & \text{otherwise} \end{cases}$
Trg	Triangle	$I(y) = \begin{cases} (y - y_{i-1})/(y_i - y_{i-1}) & y_{i-1} \le y \le y_i \\ (y_{i+1} - y)/(y_{i+1} - y_i) & y_i \le y \le y_{i+1} \end{cases}$
STrg	Sinusoidal Triangle	$I(y) = \begin{cases} \sin(k(y - y_{i-1}))/\sin(k(y_i - y_{i-1})) & y_{i-1} \le y \le y_i \\ \sin(k(y_{i+1} - y))/\sin(k(y_{i+1} - y_i)) & y_i \le y \le y_{i+1} \end{cases}$
TTS	Three Term Sinusoid	$I(y) = A_i + B_i \sin(k(y - y_i)) + C_i(\cos(k(y - y_i) - 1)$ where $y_i \le y \le y_{i+1}$
Sp1	Spline – degree 1	$I(y) = a_i + b_i(y - y_i) \quad y_i \le y \le y_{i+1}$
Sp2	Spline – degree 2	$I(y) = a_i + b_i(y - y_i) + c_i(y - y_i)^2 \quad y_i \le y \le y_{i+1}$
Sp3	Spline – degree 3	$I(y) = a_i + b_i(y - y_i) + c_i(y - y_i)^2 + d_i(y - y_i)^3$ where $y_i \le y \le y_{i+1}$
Sp4	Spline – degree 4	$I(y) = a_i + b_i(y - y_i) + c_i(y - y_i)^2 + d_i(y - y_i)^3 + e_i(y - y_i)^4$ where $y_i \le y \le y_{i+1}$

distance from the center of a dipole of length 2h and Δy_0 and Δy_{n+1} are the segment dimensions at the ends of the dipole. These definitions can be differentiated to whatever level is required by the continuity conditions of the model under investigation.

Splines. Mathematical splines were first discussed in 1946 [11]. An interesting history of their early development is reported by Ahlberg, et al [12]. A spline is a low order polynomial, such as $I = a_i + b_i y + c_i y^2 + d_i y^3 + e_i y^4 \ldots\ldots$ where the subscript i refers to spline segment i, and $1 \le i \le n$. The degree of the spline, q, is defined by the highest power of y employed. The use of splines specifically in EM work was first proposed by Davies [13], but surprisingly splines have not been widely adopted. Several reasons are possible.

The literature on splines almost all assumes the use of a square matrix and interpolation of data points. In order to satisfy these two requirements it is necessary to introduce constraints, so-called auxiliary conditions, that may not be relevant/appropriate for the EM problem at hand. In this study, and by intent, the matrix is not square, the splines are not interpolant and no artificial auxiliary conditions are enforced. With these constraints removed the use of splines becomes very attractive

Continuity between adjacent splines is enforced up to a level of $\frac{d^{q-1}I}{dy^{q-1}}$. These continuity conditions can be enforced as constraints, segment by segment, in the least-squares estimation process. This approach has two short-comings. First, it creates a high degree of redundancy and the results are not very stable. Second, the resulting matrix can become very large with $N.(q+1)$ unknowns. It is better to eliminate the majority of redundancies by direct elimination. When this is done, the result is a reduction in unknowns to $(N+q)$. For example, in the case of a quartic spline after elimination the defining coefficients are $a_1, b_1, c_1, d_1, e_1 \ldots\ldots\ldots e_N, a_{N+1}$. A similar approach was taken by Mahr [14] when considering the current on micro-strip antennas. The coefficients

a_1 and a_{N+1} are the amplitudes at the two ends of the spline sequence. The coefficient a_{N+1} is expressed in terms of $a_1, b_1, c_1, d_1, e_1, ...e_N$ and then set as a constraint in the least-squares estimation. When $q > 1$ it is necessary to match b_1 and b_{N+1}, c_1 and c_{N+1}, etc. Again, this is done using constraints.

The eight sub-domain functions in this study are identified in Table I. The three term sinusoid, TTS, is also treated as a spline. The formulation used here is more convenient than the version normally seen. In particular, it is only A_i that represents the amplitude at the beginning of the segment. It should be remembered that use of LSE provides not only a solution but also an estimate of the error in that solution. This is the so-called sum-of-the-squares of the residuals, or SSR. In this work log10(SSR) is reported as Err-sq.

Numerical Findings.

Cylinder. The current around the surface of an infinite, perfectly conducting cylinder is examined. Such a structure is included to illustrate what is possible when one examines a surface that does not include any discontinuities/edges. The excitation, $E_z^{inc}(\phi)$, is a normally incident TM wave. The current density is defined by:

$$\frac{k\eta}{4}\int J_z(\phi')H_0^{(2)}(kf(R,\phi,\phi'))d\phi' = E_z^{inc}(R,\phi,\phi^{inc})$$

This hypothetical structure is of interest because it has an analytical solution [p233, 15]. Hence a value for Y_∞ can be calculated independently. For a circumference of one wavelength, the magnitude of this value is 6.2366087D-03, for $\phi = \phi^{inc} = \pi$.

When applying the various expansion functions it is important to require that the amplitude, slopes and any higher derivatives where appropriate, be the same at +/-pi. This is achieved by applying the constraints $a_1 = a_{N+1}$, $b_1 = b_{N+1}$, etc.

As mentioned earlier, a check on the error model is to plot Err-sq against log10(n). Such plots are shown in Figure 1a and support the notion that the error does indeed decay as some power of n. Plots of log10(|I – Iref|) against log10(n) for each of the eight expansion functions are provided in Figure 1b. These results again provide credence for the use of the error model of equations (1) – (3) and provide another set of estimates for values of α for selected functions.

Half-wave dipole. As in the companion paper, Hallen's equation is used to evaluate the current on a dipole. The study is here restricted to dimensions of 2h=0.5 and a=0.007 wavelengths. . In this situation, the amplitudes of the first and last expansion functions need to be matched with the amplitudes of the edge mode segments. In the case of the triangle and sinusoidal triangle functions this is straightforward. In the case of splines additional work is needed. In the case of a_1 matching is trivial. In the case of a_{N+1} this coefficient is expressed in terms of $a_1, b_1, c_1, d_1, e_1, ...e_N$ in the quartic example and then set as a constraint in the least-squares estimation. When $q > 1$ it is also necessary to match b_1 and b_{N+1}, c_1 and c_{N+1}, etc., as required with the corresponding edge mode conditions. Again, this is done using constraints.

Half-wave dipole excited by a plane wave. The plots of Err-sq versus log10(n), shown in Figure 2a, again support the proposition that the error decays in proportion to some power of n. However, when compared with the corresponding results for the infinite cylinder two significant differences are apparent. First, the lowest level of error is only around –18.0 compared with a corresponding number of –23.0 in Figure 1a. [By comparison, the Err-sq value for an entire-domain function, the Chebyshev series of second kind, was also around –23.0]. Second, the results are much more closely grouped indicating much less difference in their relative performances in this application compared with the findings around the infinite cylinder.

The plots n Figure 2b show the convergence of log10(|I – Iref|) versus log10(n) which again support the proposition of equation (2). The lines with extreme values of slope are indicated. Again, we see values of α that are much lower than observed on the infinite cylinder. The value of Iref was 3.4196 as derived from use of entire

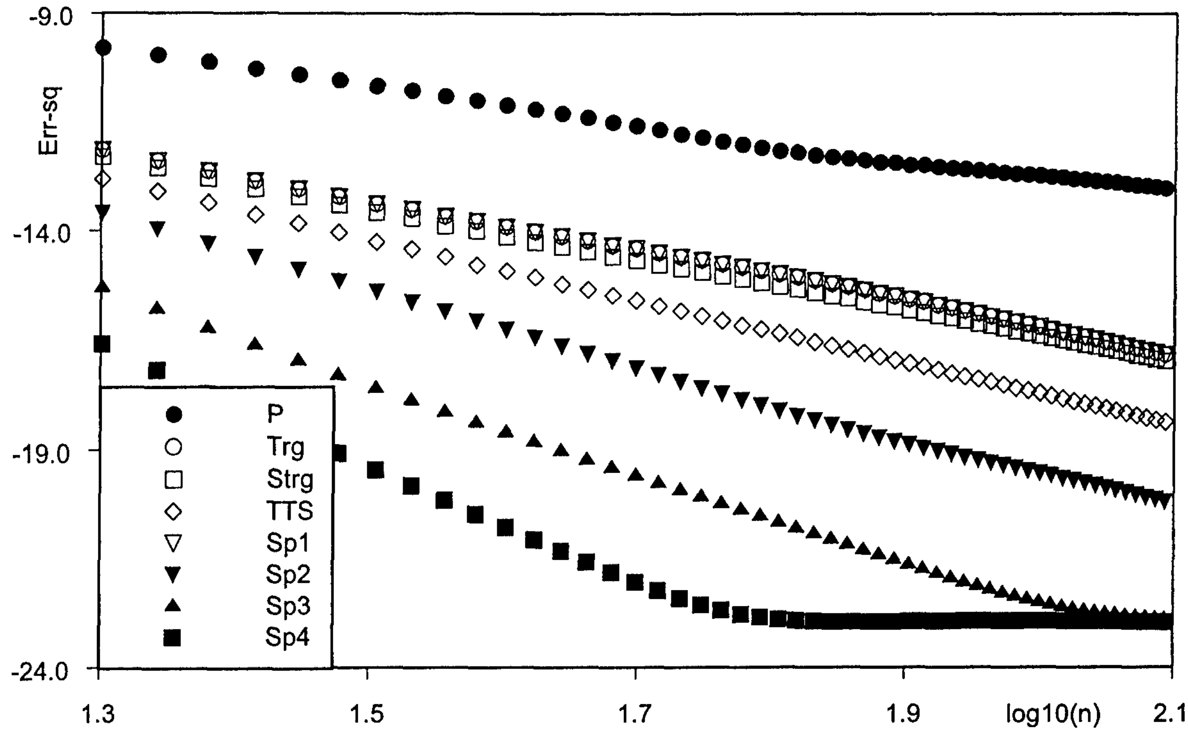

Figure 1a. Plots of Err-sq versus log10(n) for eight expansion functions, using uniform cell sizes on the surface of an infinite cylinder. Sp1, Trg and Strg almost coincide.

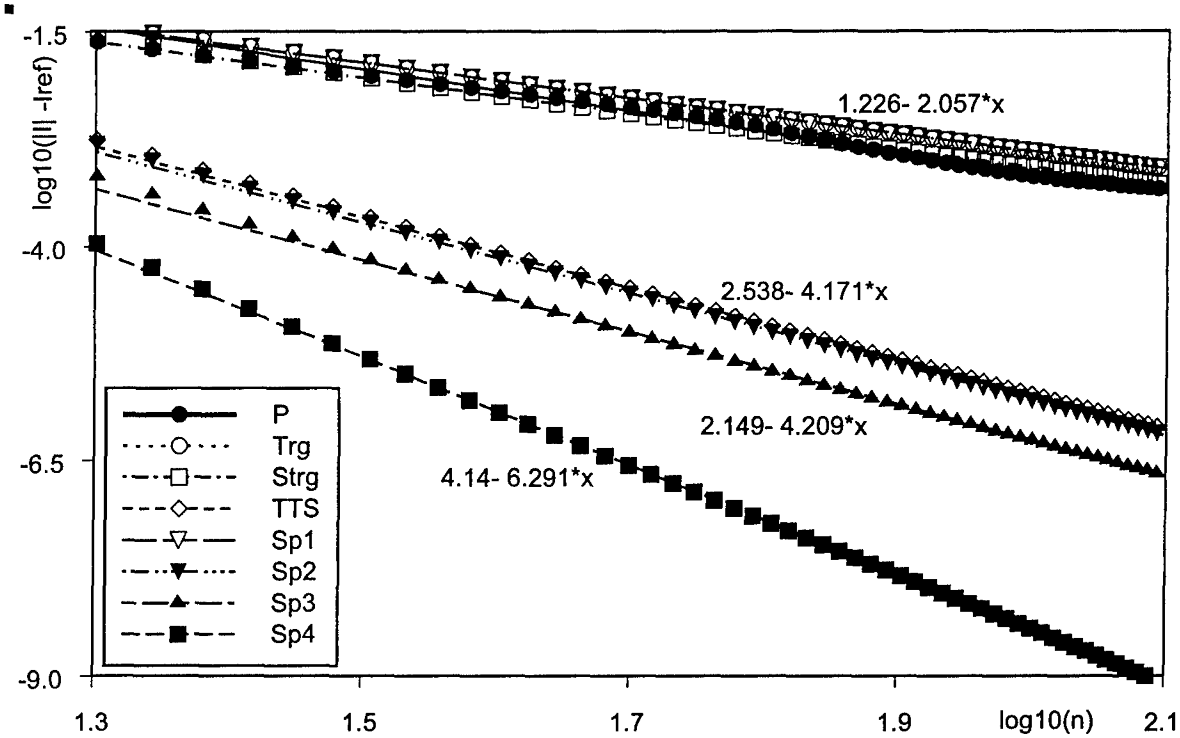

Figure 1b. Plots of log10(|I| - Iref) versus log10(n) for eight expansion functions, using uniform cell sizes on an infinite cylinder. The regression values are associated with the four splines.

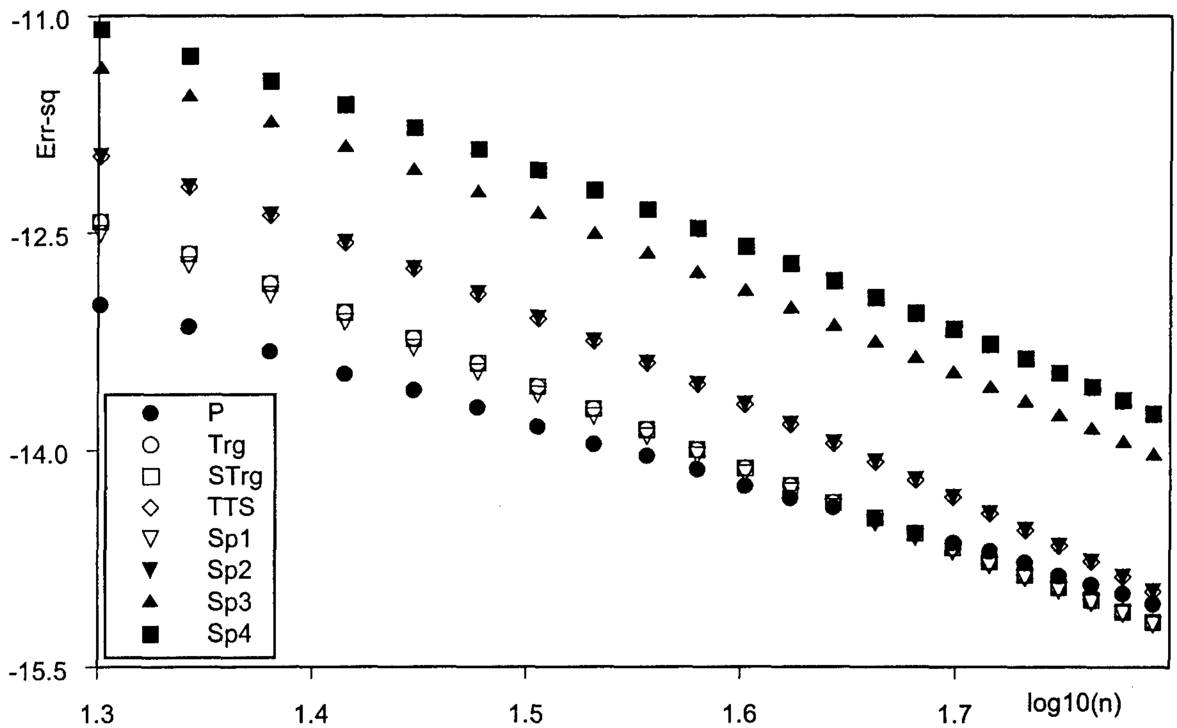

Figure 2a. Plots of Errsq versus log10(n) on the surface of a cylindrical dipole excited by a plane wave, using uniform cell sizes and all end constraints.

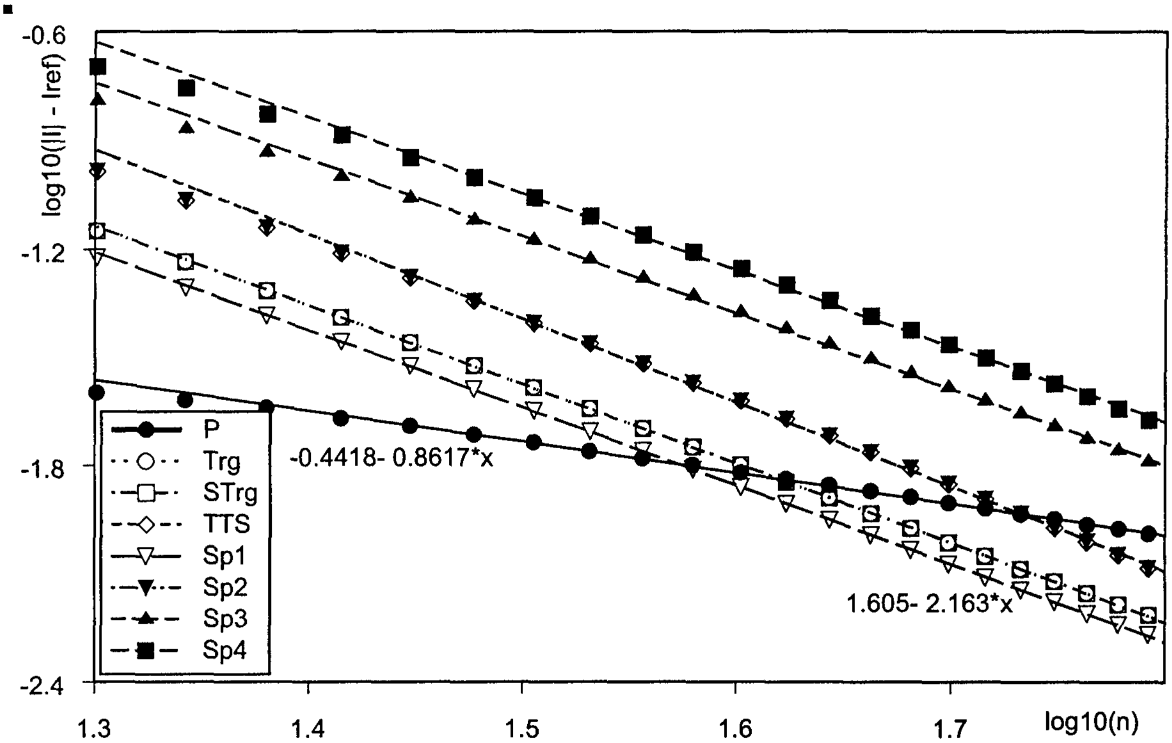

Figure 2b. Plots of log10(|I| - Iref) versus log10(n) at the center of a cylindrical dipole excited by a plane wave, using uniform cell sizes and all end constraints.

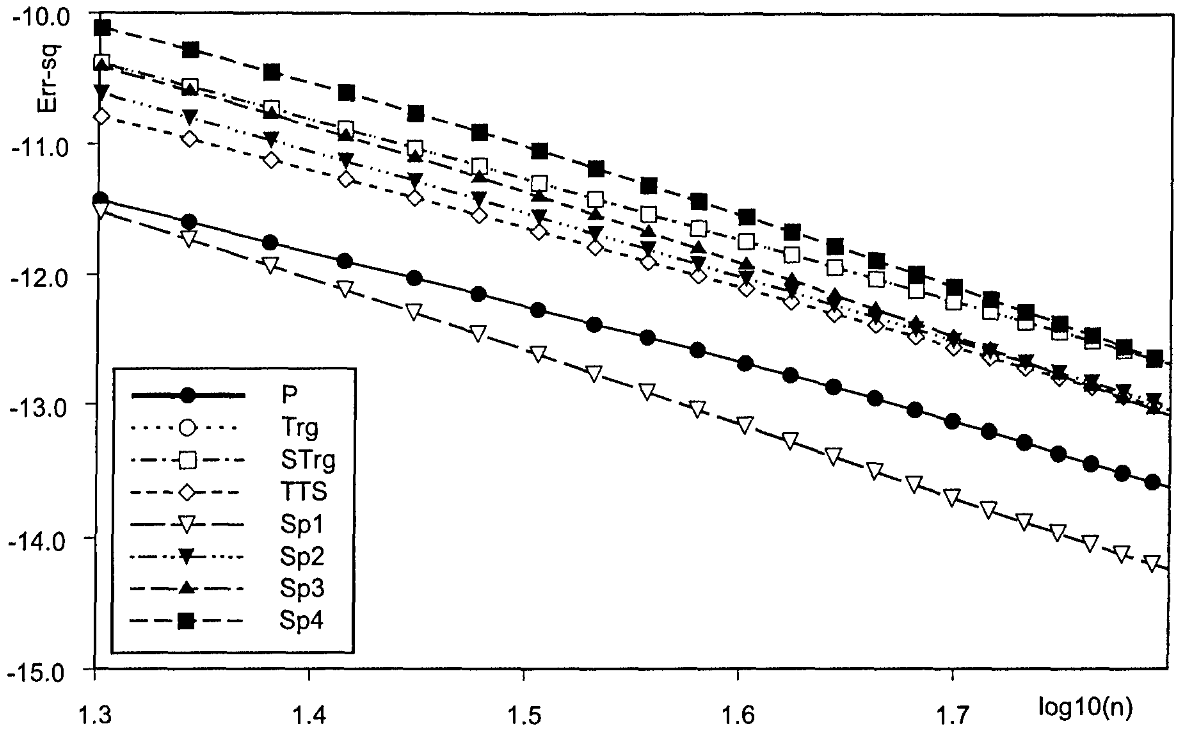

Figure 3a. Plots of Err-sq versus log10(n) for eight expansion functions on a cylindrical dipole excited by a magnetic frill, using uniform cell sizes and all end constraints.

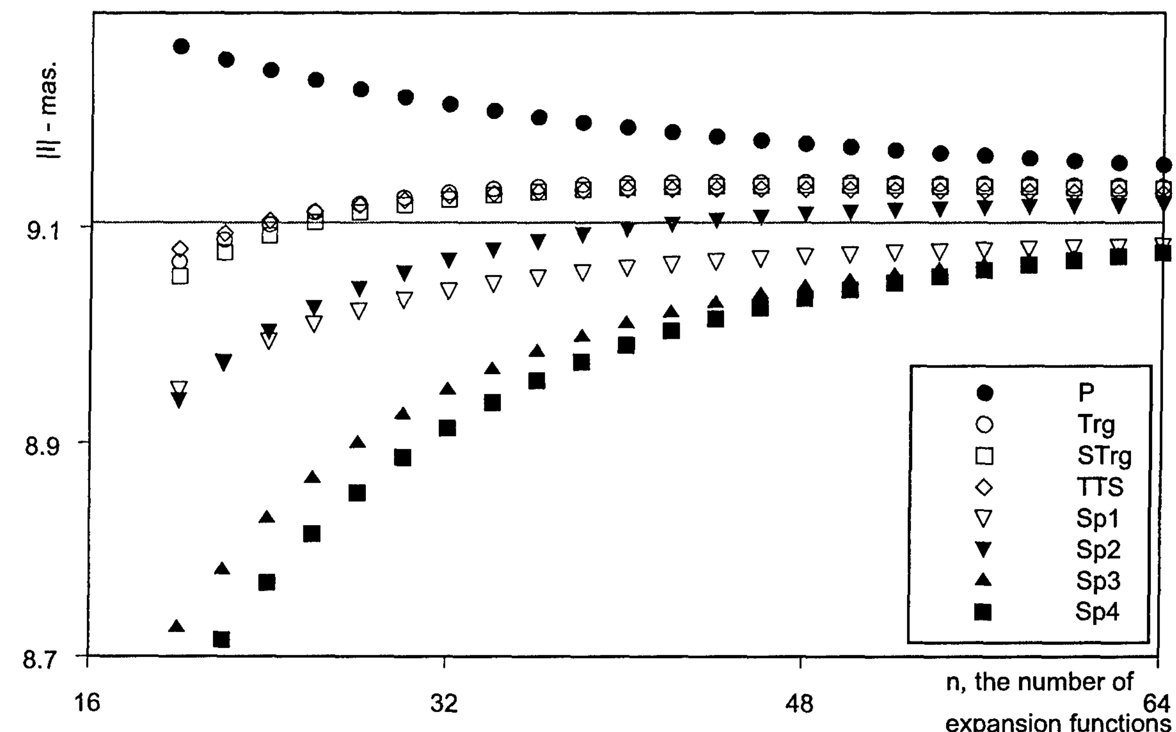

Figure 3b. Plots of |I| versus n for eight expansion functions on the surface of a dipole excited by a magnetic frill, using uniform cell sizes and all end constraints.

domain functions, and the linearity of the plots appears to support its use here. However, no statements of statistical significance can be made.

Half-wave dipole excited by a magnetic frill [16]. As illustrated in Figure 3a, the plots of Err-sq show that, with the exception of the linear spline, there is an even tighter grouping of the results than was observed for plane wave excitation.

The behavior of the magnitude of the current at the center of the dipole, |I|, is reported in Figure 3b. This suggests that all functions are converging toward the same final value. However, the model of equation (1) no longer holds for local convergence, at least in the range of study, as we see that half of the curves cross over the asymptotic value – something not predicted by (1). The study of entire domain functions, using a Fourier series in conjunction with a model for the current on an infinite dipole, provided a value of |I|=9.104 for the magnitude of the current at the center of the dipole [2h=0.50, a=0.007]. This value agrees visually with the apparent asymptotic value to which all series appear to be converging. However, this is not strong enough to make a statement concerning the magnitude of the asymptotic value.

Conclusions.

As a result of this work several findings are noted.

- Results for an infinite cylinder are generally in line with expectations. The performance of the quartic spline in this context is particularly impressive – being comparable with the entire domain cosine series results.
- In the case of the dipole excited by a plane wave, results are disappointing. The functions of higher degree did not perform as well as they did on the infinite cylinder. Figure 2b indicates that the pulse performs the worst and that there is little to separate the performance of the rest. Based on the results of Figure 2a, the two triangle series provide the least global error.
- In the case of the dipole excited by a magnetic frill, results are even more disappointing. No meaningful value can be calculated reliably. This is in contrast to the entire domain solution where it is possible to incorporate a special function to account for the specific effects of the magnetic frill and thereby achieve convergence.
- The concept that higher order splines would automatically lead to faster convergence is not generally supported. The surface of the infinite cylinder has no discontinuities and convergence increases with the order of the spline. The dipole excited by a plane wave has one discontinuity and mediocre convergence. The dipole excited by a magnetic frill has two discontinuities and convergence is totally unacceptable for any of the expansion functions.
- This result is probably not surprising to those familiar with spline theory. For example, deBoor [p22, 7] in one chapter has a heading "Uniform spacing of data can have bad consequences". To overcome this problem deBoor proposed the use of "expanded Chebyshev" points. However, deBoor also shows that, in certain situations, one can calculate the location of breakpoints to achieve the optimal convergence rate. This approach could form the basis of future research.

As a closing comment, it is this author's opinion that when it is possible to use entire domain expansion functions then they should be used. The improvements expected due to the flexibility of higher order sub-domain expansion functions cannot be realized, at least with uniform spacing.

Acknowledgements. The author thanks Professor Andrew Peterson for his encouragement and useful suggestions in connection with the preparation of this, and the companion, paper.

References.

1. M. M. Bibby, "A Comparative Study of Expansion Functions Using the Boundary Residual Method – Part I: Entire-Domain Functions", Journal of the Applied Computational Electromagnetics Society, this issue, 2002.
2. C. M. Butler & D. R. Wilton, "Analysis of Various Numerical Techniques Applied to Thin Wire Scatterers", IEEE Trans. on Ant. & Prop., pp.534-540, July, 1975.
3. K. J. Bunch & R. W. Grow, "Numeric Aspects of the Boundary Residual Method", Int. J. of Numerical Modeling: Electronic Networks, Vol. 3, pp.57-71, 1990.
4. K. J. Bunch & R. W. Grow, "On the Convergence of the Method Moments, the Boundary Residual Method, and the Point-Matching Method with a Rigorously Convergent Formulation of the Point-Matching

Method", Journal of the Applied Computational Society, Vol. 8, No 2, 1993, pp. 188-202

5. G. H. Golub & C. F. Van Loan, Matrix Computations, The Johns Hopkins Univeristy Press, 1989
6. C. L. Lawson & R. J. Hanson, Solving Least-Squares Problems, Prentice-Hall Inc., 1974
7. LAPACK 3.0, available from www.netlib.org/toms/
8. C. de Boor, A Practical Guide to Splines, Springer-Verlag, 1978
9. H. Ikuno & K. Yasuura, "Improved Point-Matching Method with Application to Scattering from a Periodic Surface", IEEE Trans. on Ant. & Prop., AP-21, 657-662, 1973
10. J. Meixner, "The behavior of Electromagnetic Fields at Edges", IEEE Trans on Ant. & Prop., Vol. 20, No 4, pp.442-446, 1972
11. I. J. Schoenberg, "Contributions to the Problem of Approximation of Equidistant Data by Analytic Functions", Quart. Appl. Math., Vol. 4, 45-99, 112-141, 1946
12. J. H. Ahlberg, E. N. Nilson & J. L. Walsh, The Theory of Splines and Their Applications, Academic Press, 1967
13. J. B. Davies, "A Least-Squares Boundary Residual Method for the Numerical Solution of Scattering Problems", IEEE Trans. M. T. & T., Vol. 21, No 2, pp.99-104, Feb. 1973
14. U. Mahr, "Cubic Splines as Expansion Functions for the Current Distribution of Microstrip Stub Antennas", IEEE Trans. on Ant. & Prop. AP-30, No 9, pp.1033-1036, Sept. 1985
15. R. F. Harrington, Time-Harmonic Electromagnetic Fields, McGraw-Hill, 1961
16. L. L. Tsai, "A Numerical Solution for the Near and Far Fields of an Annular Ring of Magnetic Current", IEEE Trans. of Ant. & Prop., Vol. 20, No 5, pp.569-576, Sept., 1972.